Phenotypic and genotypic characterization and compatibility among genotypes to select elite clones of cupuassu

Rafael Moysés ALVES¹*, Carlos Rogério de Sousa SILVA², Paulo Sérgio Bevilaqua de ALBUQUERQUE², Vinicius Silva dos SANTOS³

¹ Embrapa Amazônia Oriental, Tv. Dr. Enéas Pinheiro S/N, CP 48, CEP 66095-100, Bairro Marco, Belém, Pará, Brasil
² Estação de Recursos Genéticos José Haroldo, Comissão Executiva do Plano de Lavoura Cacaueira, BR 316, km 17, CP 46, CEP 67105-970, Marituba, Pará, Brasil
³ Universidade Federal de Viçosa (UFV), Avenida Peter Henry Rolfs, s/n, Campus Universitário, CEP 36570-900, Viçosa, Minas Gerais, Brasil

* Corresponding author: rafael-moyses.alves@embrapa.br

ABSTRACT
The expansion of the genetic base of cultivated materials is an ongoing activity of the cupuassu (Theobroma grandiflorum) breeding program. However, the parents involved need to be genotypically and phenotypically characterized to ensure compatibility of crossings, as well as to assist in the selection of more promising individuals for hybridization. This study aimed to identify and select T. grandiflorum clones that are compatible and genetically divergent using tools such as the estimates of genotypic, phenotypic, and combined distances, as well as the compatibility rates among clones. The genetic distance analysis of the clones was performed with 14 heterologous microsatellite primers of cocoa (Theobroma cacao) that amplify the DNA of cupuassu. Phenotypic characterization was based on 14 variables related to fruit production. The joint dissimilarity matrix was obtained by means of the sum of the phenotypic and molecular dissimilarity matrices. The intra- and inter-clonal compatibility was estimated through controlled crossings. A low correlation was noted between the dissimilarity matrices based on the molecular and agronomic data. As for compatibility, all clones were self-incompatible, with different compatibility rates when crossed. The compatibility index was strongly influenced by the degree of relationship of the clones. It was possible to identify and select the most promising sets of cupuassu clones to be used in breeding programs, despite their genetic relationship.

KEYWORDS: microsatellites, tropical fruit, diversity, plant breeding, Theobroma grandiflorum

Caracterização fenotípica, genotípica e compatibilidade entre genótipos para seleção de clones elite de cupuaçuzeiro

RESUMO
A ampliação da base genética dos materiais de cultivo é trabalho contínuo do programa de melhoramento genético do cupuaçuzeiro (Theobroma grandiflorum). Há necessidade, entretanto, que os parentais envolvidos estejam caracterizados fenotipicamente e genotipicamente, para auxiliar na escolha dos indivíduos que serão hibridizados, bem como para garantir os mecanismos de rastreabilidade e proteção das cultivares. Este estudo teve como objetivo identificar e selecionar clones de cupuaçuzeiro inter-compatíveis e geneticamente divergentes, utilizando como ferramentas as distâncias genotípica, fenotípica e combinada mistas, assim como as taxas de compatibilidade entre os clones. Estimativas das distâncias genéticas entre os clones foram realizadas com base em 14 iniciadores microssatélites heterólogos de cacaueiro (Theobroma cacao) que amplificam o DNA do cupuaçuzeiro. Para a caracterização fenotípica foram empregadas 14 variáveis relacionadas à produção de frutos. A matriz de dissimilaridade conjunta foi obtida por meio da soma das matrizes de dissimilaridade fenotípica e molecular. A compatibilidade intra e inter-clonal foi estimada através de cruzamentos controlados. Houve uma baixa correlação entre as matrizes de dissimilaridade com base nos dados moleculares e agronômicos. Quanto à compatibilidade, todos os clones foram auto-incompatíveis, contudo, compatíveis entre si, com diferentes taxas. O índice de compatibilidade foi fortemente influenciado pelo grau de relacionamento dos clones. Foi possível identificar e selecionar os conjuntos de clones de cupuaçuzeiro mais promissóres para ser usados em melhoramento genético, apesar da ocorrência de relação genética entre eles.

PALAVRAS-CHAVE: microssatélites, fruteira tropical, diversidade, melhoramento de plantas, Theobroma grandiflorum
INTRODUCTION

Cupuassu (*Theobroma grandiflorum* [Willd. ex Spreng.] Schum.) is a plant that is native to the Amazon and of great importance to family farmers in northern Brazil (Alves et al. 2007). In general, cupuassu plantations have low productivity due to the unselected planting material employed, with high variability and susceptibility to diseases (Alves et al. 2009). Among the diseases that occur in the cupuassu, the witch’s broom, caused by the fungus *Moniliophthora perniciosa* (Stahel: Aime & Phillips-Mora), is the most important and, when left uncontrolled, the commercial planting becomes infeasible (Alves et al. 2009). The cupuassu genetic breeding programs conducted in the Brazilian Amazon (Alves et al. 2010; Souza et al. 2002) focus on obtaining plants that are resistant to witch’s broom, and with a high production of pulp and seeds (Regazzi et al. 2002; Machado et al. 2002).

In 2012, the BRS Carimbó variety was released, which is propagated by seeds with a genetic structure of improved first-cycle population, and obtained by the crossing of 16 parents (Alves et al. 2010). Further assessments revealed that these 16 parental clones possessed exceptional characteristics of yield and resistance to witch’s broom. Thus, a clonal variety selected from a subset of these 16 clones should show a potential of high genetic gain. However, the fact that some of the 16 parental clones are related (half and full sibs) makes it necessary to employ an effective methodology that allows the identification of the subsets of more compatible and divergent clones.

The breeding programs and conservation of tree species generally use predictive methods for the analysis of genetic divergence, assessing morphological, physiological, chemical, and molecular characteristics, among others, which are usually quantified by a dissimilarity measure. However, in studies of genetic diversity different types of data are analyzed separately. One reason for this is the lack of methodologies that allow the analysis of different types of variables based on a single measure of dissimilarity (Franco et al. 1997). In order to resolve this issue, Cruz et al. (2011) proposed three different methodologies for the joint analysis of molecular and phenotypic data: (a) the sum of the standardized distances of individual matrices, (b) the transformation of quantitative variables in multicategories, and (c) the use of the distance matrix proposed by Gower (1971) by adding the phenotypic matrix with the genotypic distance matrix.

In cocoa (*Theobroma cacao* L.), the evaluation and characterization of the accessions collected and maintained in germplasm banks have been routinely practiced. Several characters have shown high variability, such as those related to the leaf, fruit, seed, flower, plant architecture, self-incompatibility, and production (Almeida et al. 2009; Santos et al. 2011; Alexandre et al. 2015).

Self-incompatibility systems are common in the genus *Theobroma*, and have been extensively studied in cocoa. In the self-incompatible MC 30 cocoa clone, the self-incompatibility mechanism operated in two stages, initially preventing pollen germination and then gametic fusion (Aneja et al. 1994). Subsequent studies demonstrated that three loci are involved in the control of cocoa fertilization (Ramos et al. 2005). Similar to *T. cacao*, *T. grandiflorum* presents pre-zygotic self-incompatibility with ovarian reaction, and the abscission of the flower before the formation of the embryo (Venturieri, 2011). The estimate of the multilocus crossing rate also suggests that *T. grandiflorum* is a perfectly allogamous species (Alves et al. 2003).

Low productivity of cupuassu is strongly influenced by reduced fertility owing to issues of pollination and self-incompatibility. The weak conversion of flowers into fruit stems from the presence of natural barriers that affect the number of pollinated flowers (Venturieri, 2011). Besides being an essentially allogamous species, with selfing rates of nearly zero, cupuassu also presents inter-clonal incompatibility, which is directly proportional to the degree of relationship between the involved parents (Falcão et al. 1999).

This study aimed to identify and select *T. grandiflorum* clones that are inter-compatible and genetically divergent using estimates of genotypic, phenotypic and combined distances, as well as the compatibility rates among clones.

MATERIALS AND METHODS

Agronomic description

The experiment for the collection of phenotypic data was conducted from 2008 to 2012 in the municipality of Tomé Açu (in the northeastern region of the state of Pará, Embrapa Eastern Amazonia; 2º25’08”S, 48º09’08” W). This region has a hot and humid climate, with an average annual temperature of 26.4°C (maximum 32.8°C, minimum 21.9°C) and an average precipitation of 2,300 mm. The soil is characterized as a Dystrophic Yellow Latosol, which is deep and of medium texture (Rodrigues et al. 2001).

The 16 parental clones (Table 1) were distributed in an experimental design of randomized complete blocks with 16 treatments (clones) and 20 replications. Each plant constituted one experimental plot. The phenotypic characterization was based on the fruit characteristics of the clones. The samples were collected from five fruits/plant/crop, measured individually as per the method described by Machado et al. (2002) in four crops between 2008 and 2012. Fourteen characteristics were used in the study: fruit longitudinal length (FLL, in mm); fruit transversal diameter (FTD, in mm); fruit rind thickness (FRT, in mm); fruit weight (FW, in g); rind weight (RW, in g); fresh seed weight (FSW, in g); dry seed weight (DSW, in g); pulp
weight (PW, in g); number of seeds per fruit (NS); number of
fruits per plant (NP); pH; total soluble solids (TSS); titratable
acidity (TTA); and total solids (TS).

Molecular characterization

To perform the genetic characterization, leaf tissues of
clones were subjected to DNA extraction in the molecular
biology laboratory of CEPLAC in Marituba, Pará. The
extraction was performed according to the protocol of
Doyle and Doyle (1990) as modified by Figueira et al.
(1997) for Theobroma spp. The molecular markers used
were microsatellites of T. cacao, which were developed by
Lanaud et al. (1999) and Pugh et al. (2004). We employed 14
heterologous primers of T. cacao microsatellites that amplify
the DNA of cupuassu, based on the selection made by Alves
et al. (2007). The reactions were prepared in a final volume
of 13 µl containing 15 ng of genomic DNA; 100 µM of each
dNTP; 0.2 µM of each primer (forward and reverse); enzyme
buffer (50 mM KCl, 10 mM Tris-HCl; pH 8.8, 0.1% Triton
X-100, 1.5 mM MgCl2), and 1 unit of Taq DNA polymerase.
The reactions were amplified in a Thermocycler GeneAmp
PCR System 9700 (Applied Biosystems; Foster City, USA),
which was initially programmed with a cycle of denaturation
at 94°C for 4 min; followed by 10 cycles of 94°C for 40 s,
with a decrease of 1°C by cycle, with an initial temperature of
59°C and final temperature of 49°C for 40 s, and, finally,
an extension at 72°C for 60 s. The amplified fragments
were separated on polyacrylamide denaturant gel at 6% and
revealed using silver nitrate, as per the method of Creste et al.
(2001). The profile of each tested material was obtained to
serve as the “fingerprint” of genetic characterization in order
to ensure the identity and traceability of the clones.

Interclone compatibility

To define the interclonal compatibility rate, we performed
controlled crossings in a partial diallel design of 16 x 17, with
10 pollinations per crossing, repeated in two seasons (2010 and
2011). In this study, the clone 186 (Codajás) (Table 1) was added
to the 16 evaluated genotypes because it is a parent of eight of
the tested materials. We performed 136 crossings in total.

Evaluations were performed at 15 days after pollination,
when the fruit formation was perfectly defined. For each
crossing, the percentage of fruit sets resulting from 10
pollinations was calculated, which allowed us to estimate the
average compatibility rate of each clone. The standard deviation
was also estimated as a measure of intraclonal variability. We
used an adaptation of the methodology suggested by Zapata
and Arroyo (1978), the index “Percentage of High Crossings”
(PHC), which assumes that compatibility rates above a cut-off
of 30% are indicative of good compatibility between the
materials involved in the crossings: PHC = (NCS x 100)/TC,
where NCS = the number of crossings of a specific clone with
a compatibility rate of >30% and TC = the total number of
crossings of the specific clone.

Statistical analysis

The genetic diversity study was based on cluster
analysis using phenotypic and genotypic data, as well
as the combination of both. For the phenotypic data, a
multicollinearity diagnostic was performed based on the
condition number (CN) proposed by Montgomery et al.
(2001), where CN < 100 indicates weak, 100 ⩽ CN ⩽ 1000
moderate to strong, and CN > 1000 severe multicollinearity.
If multicollinearity is observed in moderate to severe levels,
the highest correlated variables should be excluded (Cruz et al.
2014). Severe multicollinearity can alter the hierarchical
clustering pattern because multicollinear variables are given
greater weight in the analysis (Corrêa et al. 2007; Cargnelutti
Filho et al. 2009). A Gower’s distance matrix was generated
for the uncorrelated agronomic traits, with distance measure
ranging from 0 to 1 (Cruz et al. 2011).

For the molecular data, genetic distances among clones
were estimated by the arithmetic complement of the
unweighted index (ACUI), which also ranges from 0 to 1
(Cruz et al. 2011). The combined distance of genotypic and
phenotypic data was obtained by the sum of the Gower and

<table>
<thead>
<tr>
<th>Clone</th>
<th>Ancestry</th>
<th>Origin of mother</th>
<th>Origin of father</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>174 x 186</td>
<td>174: Coari – AM</td>
<td>186: Codajás – AM</td>
</tr>
<tr>
<td>42</td>
<td>186 x 434</td>
<td>186: Codajás – AM</td>
<td>434: Muná – PA</td>
</tr>
<tr>
<td>44</td>
<td>186 x 434</td>
<td>186: Codajás – AM</td>
<td>434: Muná – PA</td>
</tr>
<tr>
<td>46</td>
<td>186 x 215</td>
<td>186: Codajás – AM</td>
<td>215: Manacapuru – AM</td>
</tr>
<tr>
<td>47</td>
<td>186 x 1074</td>
<td>186: Codajás – AM</td>
<td>1074: Itacoatiara – AM</td>
</tr>
<tr>
<td>48</td>
<td>186 x 1074</td>
<td>186: Codajás – AM</td>
<td>1074: Itacoatiara – AM</td>
</tr>
<tr>
<td>51</td>
<td>215 x 624</td>
<td>215: Manacapuru – AM</td>
<td>624: Santarém – PA</td>
</tr>
<tr>
<td>56</td>
<td>186 x 1074</td>
<td>186: Codajás – AM</td>
<td>1074: Itacoatiara – AM</td>
</tr>
<tr>
<td>57</td>
<td>186 x 513</td>
<td>186: Codajás – AM</td>
<td>513: Oiapoque – AP</td>
</tr>
<tr>
<td>61</td>
<td>220 x 228</td>
<td>220: Manacapuru – AM</td>
<td>228: Manaus – AM</td>
</tr>
<tr>
<td>62</td>
<td>220 x 185</td>
<td>220: Manacapuru – AM</td>
<td>185: Codajás – AM</td>
</tr>
<tr>
<td>63</td>
<td>174 x 248</td>
<td>174: Coari – AM</td>
<td>248: Itacoatiara – AM</td>
</tr>
<tr>
<td>64</td>
<td>220 x 185</td>
<td>220: Manacapuru – AM</td>
<td>185: Codajás – AM</td>
</tr>
<tr>
<td>174</td>
<td>Original collection</td>
<td>174: Coari – AM</td>
<td>-</td>
</tr>
<tr>
<td>215</td>
<td>Original collection</td>
<td>215: Manacapuru – AM</td>
<td>-</td>
</tr>
<tr>
<td>1074</td>
<td>Original collection</td>
<td>1074: Itacoatiara - AM</td>
<td>-</td>
</tr>
<tr>
<td>186</td>
<td>Original collection</td>
<td>186: Codajás – AM</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1. Ancestry and origin of T. grandiflorum clones evaluated in this study. AM indicates origin in the state of Amazonas, Brazil; PA indicates origin in the state of Pará, Brazil.
ACUI distance matrices, as described by Cruz et al. (2011) and Alves et al. (2013).

The three dissimilarity matrices were used in a hierarchical clustering analysis based on the complete linkage (farthest neighbor) method. The cut-off point for cluster separation in the dendrogram was set according to Mojena’s criterion (Mojena, 1977), which aims to select the number of groups in step j that, first, satisfies the following inequality: $a_j > a + k s_j$, where $a_j$ is the distance value of the fusion level corresponding to steps j, (for j = 1, 2, ..., g-1), where $a$ is the mean of $a_j$, given by:

$$a = \frac{1}{g} \sum_{j=1}^{g} a_j$$

and $s$ is the standard deviation of $a_j$, given by:

$$s = \sqrt{\frac{1}{g-1} \sum_{j=1}^{g} (a_j - a)^2}$$

and $k$ is a constant equal to 1.25, as suggested by Milligan and Cooper (1985). All analyses were conducted using the Genes software (Cruz, 2013).

For the selection process, the promising set of clones were identified through most of the compatible clones for each reference clone, considering the PHC index. From these groupings, we sequentially removed clones with very close genotypic, phenotypic, and combined distances. The remaining clones were assumed to be most favorable genotypes for the production of hybrids with the reference clone or polyclonal varieties.

**RESULTS**

We obtained 41 fragments from the 14 microsatellite loci evaluated, an average of 2.9 alleles/locus (Table 2). The most polymorphic loci were mTcCIR 7, mTcCIR 25, mTcCIR99, and mTcCIR 135.

The hierarchical clustering of clones considering only molecular aspects formed three groups (Figure 1). In Group I, a prevalence of the ancestral clone 220 from the Manacapuru region (Amazonas State) was noted. In Group II, all clones were hybrids of clone 186 (Codajás-AM). Group III was composed of clones originating from Iraçatiara and Codajás (Amazonas State). The average distance based on ACUI was 0.455. The smallest distance observed between clones of the same genetic origin was 0.192 for clones 47 and 48 (both originated from the crossing of 186 and 1074). The greatest distance was 0.727, as observed between clones without geographic or genetic relationships.

Severe multicollinearity among phenotypic variables was noted. After exclusion of the variables rind thickness (RT), fruit weight (FW), and dry seed weight (DSW), CN was reduced to 76.51, indicating low multicollinearity. Fruit weight is an important variable, but it is highly multicollinear as it represents the sum of the variables pulp weight (PW), rind weight (RW), and fresh seed weight (FSW). The clustering of clones based on phenotypic data (Figure 2) showed the formation of two groups: Group I consisted of 7 clones, and mTcCIR 286.

**Table 2.** SSR primers used in the study, location in the linkage group, estimated size (bp), and the number of alleles in the analysis of 16 clones of *T. grandiflorum* at Embrapa Eastern Amazonia, Pará State, Brazil.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Linkage group</th>
<th>Size of cloned allele (bp)</th>
<th>Estimated size (bp)</th>
<th>Number of alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>mTcCIR 7</td>
<td>VII</td>
<td>160</td>
<td>147, 149, 150, 152</td>
<td>4</td>
</tr>
<tr>
<td>mTcCIR 9</td>
<td>VII</td>
<td>274</td>
<td>295, 310</td>
<td>2</td>
</tr>
<tr>
<td>mTcCIR 17</td>
<td>IV</td>
<td>271</td>
<td>280, 292</td>
<td>2</td>
</tr>
<tr>
<td>mTcCIR 25</td>
<td>VI</td>
<td>153</td>
<td>134, 135, 147, 170</td>
<td>4</td>
</tr>
<tr>
<td>mTcCIR 26</td>
<td>VIII</td>
<td>298</td>
<td>250, 265, 270</td>
<td>3</td>
</tr>
<tr>
<td>mTcCIR 54</td>
<td>I</td>
<td>165</td>
<td>149, 150, 151</td>
<td>3</td>
</tr>
<tr>
<td>mTcCIR 57</td>
<td>IV</td>
<td>253</td>
<td>260, 268, 270</td>
<td>3</td>
</tr>
<tr>
<td>mTcCIR 99</td>
<td>VIII</td>
<td>249</td>
<td>270, 280, 288, 300</td>
<td>4</td>
</tr>
<tr>
<td>mTcCIR 124</td>
<td>IX</td>
<td>131</td>
<td>155, 162</td>
<td>2</td>
</tr>
<tr>
<td>mTcCIR 135</td>
<td>III</td>
<td>246</td>
<td>254, 260, 270, 280</td>
<td>4</td>
</tr>
<tr>
<td>mTcCIR 162</td>
<td>III</td>
<td>162</td>
<td>190, 200, 204</td>
<td>3</td>
</tr>
<tr>
<td>mTcCIR 220</td>
<td>X</td>
<td>201</td>
<td>225, 230</td>
<td>2</td>
</tr>
<tr>
<td>mTcCIR 270</td>
<td>I</td>
<td>224</td>
<td>220, 222</td>
<td>2</td>
</tr>
<tr>
<td>mTcCIR 286</td>
<td>I</td>
<td>119</td>
<td>290, 300, 304</td>
<td>3</td>
</tr>
</tbody>
</table>

**Figure 1.** Dendrogram resulting from the hierarchical clustering analysis of genotypic characters of cupuassu (*T. grandiflorum*) clones using the complete linkage (farthest neighbor) method and the measure of dissimilarity of unweighted arithmetic complement. The data are from cupuassu clones that originated in the states of Amazonas and Pará, Brazilian Amazon and were cultivated experimentally at Embrapa Eastern Amazonia, Pará, Brazil.

**Figure 2.** Dendrogram resulting from the hierarchical clustering analysis based on 11 phenotypic traits of cupuassu (*T. grandiflorum*) clones using the complete linkage (farthest neighbor) method and Gower’s distance matrix. The data are from the cupuassu clones that originated in the states of Amazonas and Pará, Brazilian Amazon and were cultivated experimentally at Embrapa Eastern Amazonia, Pará, Brazil.
including five descendants of clone 186 (Codajás, Amazonas), as was also observed in the genotypic grouping. Group II was formed by primary clones 174 and 1074, along with their descendants. Three progenies of clone 220 were also allocated to this group in accordance with the results of genotypic data (Figures 1 and 2).

The correlation between the matrices of genotypic and phenotypic distances was 0.02. The correlations between the matrix of combined distances (sum of the individual matrices) and the genotypic and phenotypic matrices were 0.80 and 0.58, respectively; both were significant by the z and Mantel tests at \( P \leq 0.01 \) with 1,000 simulations. The clustering based on the combined distance matrix formed three groups (Figure 3). With the exception of clone 61, all clones allocated to Group I were also assigned to the same group in the clustering based on genotypic information, but to different groupings in the phenotypic data clustering. Group II was composed of four clones, which were also allocated to the same group in the genotypic clustering; and, with the exception of clone 215, the remaining clones also belonged to the same group based on the phenotypic data. Group III was formed by the remaining six clones, of which 4 belonged to the same group based on genotypic information.

The compatibility rate between full sib individuals and backcrossings was low (generally around 20%), 50% of the crossings between half sibs showed high values of compatibility, while only 19% of the crossings between unrelated individuals showed low compatibility (Table 3). Clones 1074, 64, 56, and 61 had PHC > 75%, while other crossings involving half sibs showed a compatibility of ≤10% (crossings 32 x 46; 32 x 57; 46 x 47; 46 x 57). Some half-sib crossings were so incompatible that they could not produce any fruit. Clone 47 was present in all the crossings, and thus may have been the cause of this extreme incompatibility. In the five crossings performed within each group of full sibs, 80% showed low compatibility values.

![Dendrogram resulting from the hierarchical clustering analysis of the matrix of combined genotypic and phenotypic dissimilarities of cupuassu (T. grandiflorum) clones using the complete linkage (farthest neighbor) method. The data are from cupuassu clones that originated in the states of Amazons and Pará, Brazilian Amazon, and were cultivated experimentally at Embrapa Eastern Amazonia, Pará, Brazil.](image)

**Table 3.** Measured compatibility matrix, average compatibility rate (ACR), standard deviation (SD), and percentage of high crossings (PHC) among the crossings made with 16 genotypes of *T. grandiflorum* (parents of the BRS Carimbó variety) in 2010 and 2011 in Belém and Tomé Açu, Pará State, Brazil.

<table>
<thead>
<tr>
<th>Parental</th>
<th>32</th>
<th>42</th>
<th>44</th>
<th>46</th>
<th>47</th>
<th>48</th>
<th>51</th>
<th>56</th>
<th>57</th>
<th>61</th>
<th>62</th>
<th>63</th>
<th>64</th>
<th>174</th>
<th>1074</th>
<th>186</th>
<th>ACR</th>
<th>SD</th>
<th>PHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>15.0</td>
<td>56.7</td>
<td>7.5</td>
<td>0.0</td>
<td>45.0</td>
<td>18.4</td>
<td>40.0</td>
<td>2.5</td>
<td>35.0</td>
<td>15.0</td>
<td>22.5</td>
<td>30.0</td>
<td>23.9</td>
<td>41.3</td>
<td>65.0</td>
<td>6.8</td>
<td>26.5</td>
<td>19.3</td>
<td>43.8</td>
</tr>
<tr>
<td>42</td>
<td>27.5</td>
<td>25.0</td>
<td>0.0</td>
<td>37.5</td>
<td>40.7</td>
<td>44.2</td>
<td>17.5</td>
<td>60.0</td>
<td>15.0</td>
<td>35.0</td>
<td>45.0</td>
<td>38.4</td>
<td>42.1</td>
<td>38.4</td>
<td>32.5</td>
<td>33.2</td>
<td>14.8</td>
<td>62.5</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>77.5</td>
<td>46.7</td>
<td>21.7</td>
<td>35.0</td>
<td>15.0</td>
<td>75.0</td>
<td>42.5</td>
<td>15.0</td>
<td>67.5</td>
<td>85.0</td>
<td>83.8</td>
<td>85.0</td>
<td>90.0</td>
<td>24.4</td>
<td>52.8</td>
<td>27.7</td>
<td>68.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>10.0</td>
<td>85.0</td>
<td>27.5</td>
<td>85.0</td>
<td>5.0</td>
<td>100.0</td>
<td>100.0</td>
<td>25.0</td>
<td>90.0</td>
<td>38.8</td>
<td>22.5</td>
<td>65.0</td>
<td>0.0</td>
<td>50.4</td>
<td>37.0</td>
<td>50.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>0.0</td>
<td>80.0</td>
<td>79.2</td>
<td>0.0</td>
<td>40.0</td>
<td>45.0</td>
<td>90.0</td>
<td>35.0</td>
<td>60.0</td>
<td>80.0</td>
<td>56.3</td>
<td>5.0</td>
<td>41.8</td>
<td>33.2</td>
<td>62.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>33.4</td>
<td>25.0</td>
<td>47.5</td>
<td>45.0</td>
<td>45.0</td>
<td>20.0</td>
<td>65.0</td>
<td>80.0</td>
<td>27.5</td>
<td>2.5</td>
<td>38.7</td>
<td>23.9</td>
<td>62.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>40.0</td>
<td>22.5</td>
<td>15.0</td>
<td>35.0</td>
<td>65.0</td>
<td>45.0</td>
<td>85.0</td>
<td>12.9</td>
<td>44.6</td>
<td>21.9</td>
<td>40.2</td>
<td>21.6</td>
<td>62.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>37.5</td>
<td>25.0</td>
<td>40.0</td>
<td>50.0</td>
<td>40.0</td>
<td>40.0</td>
<td>70.0</td>
<td>31.3</td>
<td>0.0</td>
<td>41.5</td>
<td>22.1</td>
<td>75.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>35.0</td>
<td>15.0</td>
<td>65.0</td>
<td>50.0</td>
<td>65.0</td>
<td>70.0</td>
<td>85.0</td>
<td>2.5</td>
<td>42.8</td>
<td>31.3</td>
<td>56.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>30.0</td>
<td>5.0</td>
<td>40.0</td>
<td>65.0</td>
<td>10.0</td>
<td>35.0</td>
<td>70.0</td>
<td>44.5</td>
<td>26.4</td>
<td>75.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>50.0</td>
<td>0.0</td>
<td>25.0</td>
<td>70.0</td>
<td>20.0</td>
<td>50.0</td>
<td>37.0</td>
<td>24.9</td>
<td>56.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>45.0</td>
<td>0.0</td>
<td>7.5</td>
<td>80.0</td>
<td>95.0</td>
<td>48.3</td>
<td>29.5</td>
<td>68.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>25.0</td>
<td>70.0</td>
<td>35.0</td>
<td>35.0</td>
<td>44.0</td>
<td>22.9</td>
<td>81.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>174</td>
<td>2.5</td>
<td>65.0</td>
<td>95.0</td>
<td>50.2</td>
<td>28.8</td>
<td>68.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>215</td>
<td>7.0</td>
<td>90.0</td>
<td>52.2</td>
<td>31.1</td>
<td>68.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1074</td>
<td>85.0</td>
<td>55.2</td>
<td>22.9</td>
<td>87.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>186</td>
<td>40.6</td>
<td>37.0</td>
<td>50.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR</td>
<td>43.7</td>
<td>26.7</td>
<td>64.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When backcrossings were performed with the recurrent parents 174 (Coari), 186 (Codajás), 215 (Manacapurú), and 1074, fruit set rates were also low. This was observed for crossings 32 x 174, 63 x 174, 46 x 215, 51 x 215, and 48 x 1074. In 80% of the 15 backcrossings performed, the compatibility rate was <30%. Clone 186 (Codajás), ancestor of 50% of the evaluated materials, showed an extremely low compatibility when backcrossed with its descendants. On the other hand, when crossed with other genetically unrelated materials, the average compatibility of clone 186 (Codajás) was 74%.

Among the 16 clones studied, PHC > 30% varied from 7 (clone 32) to 14 (clone 1074) in crossings, generally with higher values for more distantly related clones. After removing genetically and morphologically related clones, and the combined genetic-morphologic relationships, each clone had 3 (clones 47 and 63) to 9 (clone 186) other clones with which it could be crossed to continue the breeding program (Table 4).

**DISCUSSION**

The 16 clones of T. grandiflorum studied here are grouped either by geographic origin (with unknown genetic relation) or by close genetic kinship. Another study of genetic diversity in cupuassu using molecular markers reported that the evaluated genotypes showed genetic divergence with a moderate level of allelic richness (Alves et al. 2013). In the present study, the number of groups formed based on the hierarchical clustering considering phenotypic and genotypic data, as well as the combination of both, varied between two and three groups. This was due to the fact that most of the clones have common ancestors, wherein half of the treatments are descendants

---

**Table 4. Groups of clones of T. grandiflorum recommended for the final selection of genotypes based on the inter-clonal compatibility rates and genotypic and phenotypic distances among the clones. Data are from 16 cupuassu clones that originated in the states of Amazonas and Pará, Brazilian Amazon and were cultivated experimentally at Embrapa Eastern Amazonia, Pará State, Brazil.**

<table>
<thead>
<tr>
<th>Clone</th>
<th>Compatibility group (PHC* ≥ 30%)</th>
<th>Related clones** (low genetic distance)</th>
<th>Related clones** (low Gower distance)</th>
<th>Related clones** (low combined distances)</th>
<th>Final Selection of Genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>44, 48, 56, 61, 64, 215, 1074</td>
<td>44, 48</td>
<td>44, 215</td>
<td>48, 61</td>
<td>32, 56, 64, 1074</td>
</tr>
<tr>
<td>42</td>
<td>48, 51, 56, 61, 63, 64, 174, 215, 1074, 186</td>
<td>48</td>
<td>51, 215</td>
<td>48, 61</td>
<td>42, 56, 63, 64, 174, 1074, 186</td>
</tr>
<tr>
<td>44</td>
<td>32, 46, 47, 51, 57, 61, 63, 64, 174, 215, 1074</td>
<td>32, 46, 47</td>
<td>32, 46, 51, 57, 215</td>
<td>51, 57, 63, 174</td>
<td>44, 61, 64, 1074</td>
</tr>
<tr>
<td>46</td>
<td>44, 48, 56, 61, 62, 64, 174, 1074</td>
<td>44, 48</td>
<td>44</td>
<td>48, 61</td>
<td>46, 56, 62, 64, 174, 1074</td>
</tr>
<tr>
<td>47</td>
<td>44, 51, 56, 61, 62, 63, 64, 174, 215, 1074</td>
<td>44</td>
<td>56, 61, 62, 63, 64, 174, 1074</td>
<td>61</td>
<td>47, 51, 215</td>
</tr>
<tr>
<td>51</td>
<td>42, 44, 47, 48, 56, 62, 63, 64, 174, 215, 1074</td>
<td>56, 63</td>
<td>42, 44</td>
<td>44, 56, 63, 174</td>
<td>51, 47, 48, 62, 64, 1074</td>
</tr>
<tr>
<td>56</td>
<td>32, 42, 46, 47, 51, 57, 62, 63, 64, 174, 215, 1074</td>
<td>51, 57, 63</td>
<td>47, 62, 63, 64, 174, 1074</td>
<td>51, 57, 63, 174</td>
<td>56, 32, 42, 46, 215</td>
</tr>
<tr>
<td>57</td>
<td>44, 48, 56, 61, 63, 64, 174, 215, 1074</td>
<td>56, 63</td>
<td>44, 215</td>
<td>44, 56, 63, 174</td>
<td>57, 48, 61, 64, 1074</td>
</tr>
<tr>
<td>61</td>
<td>32, 42, 44, 46, 47, 48, 57, 62, 64, 174, 215, 1074, 186</td>
<td>62, 64, 174</td>
<td>47, 48, 62, 64, 174, 1074</td>
<td>32, 42, 46, 47, 48</td>
<td>61, 44, 57, 186</td>
</tr>
<tr>
<td>63</td>
<td>42, 44, 47, 51, 56, 57, 62, 64, 1074, 186</td>
<td>51, 56, 57</td>
<td>47, 48, 56, 62, 64, 1074, 186</td>
<td>44, 51, 56, 57</td>
<td>63, 42, 186</td>
</tr>
<tr>
<td>64</td>
<td>32, 42, 44, 46, 47, 51, 56, 57, 61, 63, 215, 1074, 186</td>
<td>61, 215, 1074</td>
<td>47, 56, 61, 137, 215, 1074</td>
<td>215, 1074</td>
<td>64, 32, 42, 44, 46, 51, 57, 186</td>
</tr>
<tr>
<td>174</td>
<td>42, 44, 46, 47, 51, 56, 57, 61, 63, 174, 215, 1074, 186</td>
<td>61, 1074</td>
<td>47, 48, 56, 1074</td>
<td>44, 51, 56, 57</td>
<td>174, 42, 186</td>
</tr>
<tr>
<td>215</td>
<td>32, 42, 44, 47, 48, 56, 57, 62, 64, 1074, 186</td>
<td>62, 64, 1074</td>
<td>32, 42, 44, 57</td>
<td>62, 64, 1074</td>
<td>215, 47, 48, 56, 186</td>
</tr>
<tr>
<td>1074</td>
<td>32, 42, 44, 47, 51, 56, 57, 61, 63, 64, 174, 215, 186</td>
<td>61, 64, 174</td>
<td>47, 56, 61, 63, 174, 215</td>
<td>64, 215</td>
<td>1074, 32, 42, 44, 46, 51, 57, 186</td>
</tr>
</tbody>
</table>

*PHC: Percentage of High Crossings (percentage of crossings with compatibility rates ≥ 30%).
** Deleted clones of the compatibility group due to their genotypic proximity to other genotypes of the groups.
*** Deleted clones of the compatibility group due to their phenotypic proximity to other genotypes of the groups.
**** Deleted clones of the compatibility group due to their phenotypic and genotypic combined proximity to other genotypes of the groups.
of clone 186. Similarly, Maia et al. (2011) obtained two groupings for 36 progenies of *T. grandiflorum* using the Tocher method (RAO, 1952) based on phenotypic data. Santos et al. (2005) found low genetic diversity among 30 cocoa clones based on the RAPD markers and genealogy information, which can also be attributed to the strong genetic relation among most of the studied materials. Studies of cupuassu diversity based on phenotypic data were performed by Araújo et al. (2002), Maia et al. (2011), and Santos et al. (2015), among others, and obtained a reasonable phenotypic divergence, with the number of groups varying between two and seven, according to several grouping methods.

The lack of correlation between the matrices of genotypic and phenotypic distances was also observed in sweet potatoes (Koussao et al. 2014) and yellow passion fruit (Reis et al. 2011). The weak association between the morphological and molecular data is due to the fact that the microsatellite markers represent only a small part of the genome and did not come from the same genomic regions that are relevant to the phenotypic expression of the variables measured. Moreover, molecular markers have been assumed to be selectively neutral, while the parts of the genome associated with phenotypic expression are subject to selection under environmental influence (Collard et al. 2005; Alves et al. 2013).

The correlations between the matrices of combined distances and based on the genotypic and phenotypic information, at high (0.80) and moderate (0.58) levels, suggest that genetic diversity analysis based on the sum of the matrices can be an important tool to cumulatively allocate genotypic and phenotypic informations, as it keeps high correlation with both the matrices (Alves et al. 2013). In addition, an average similarity >50% was observed, indicating that the variability among cupuassu clones is limited.

The compatibility between related individuals was low, while that among unrelated individuals was 81%. The results are a further demonstration that the genetic distance between parents favors fruit set. The average overall compatibility among clones presented in this research is in agreement with the results obtained by Silva et al. (2004) and Alves et al. (2007).

The detected genetic diversity will be useful to the cupuassu breeding program, and the genotypic evaluations in relation to the fruit yield and resistance to *M. perniciosa* should be intensified so that some of these clones can be recommended to the productive sector in the future. However, this detected diversity will not be sufficient to support, in the long term, the cupuassu breeding program. With increasing acreage must occur the emergence of new pathogens. There is also a need to improve the productivity and quality of pulp and seeds to meet the future demands of the agribusiness sector. For cocoa, similar results as ours have been reported by Faleiro et al. (2001), who studied the clonal varieties recommended to the cocoa producers.

For the continuity of the cupuassu breeding program, these clones could be used in controlled crossings; however, some care must be taken in relation to the restricted ancestry of the parents of these clones. Table 4 reports that the crossings would be more feasible when taking into account the phenotypic and genotypic distances, especially the combination of both in the identification of combinations that could possibly lead to a greater heterotic vigor. Excluding the clones that have been already released (174, 186, and 215), the group of seven clones, from 32, 42, 44, 46, 51, 57, and 64, seem to be the most promising for recommendation to the breeding programs. The remaining clones 47, 48, 56, 61, 63, and 1074 were discarded.

**CONCLUSIONS**

We found a reasonable divergence among the 16 studied clones of cupuassu, *T. grandiflorum*, despite their being genetically related. Based on the compatibility rates, and on the genotypic, phenotypic, and combined genotypic/phenotypic distances among the clones, we characterized and selected the most promising sets of genotypes to guide the breeding programs of *T. grandiflorum*.

**ACKNOWLEDGMENTS**

We thank the Embrapa Eastern Amazonia support staff for their assistance during the experiments. This research was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). VSS thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for his doctoral scholarship.

**REFERENCES**


Alves, R.M.; Artero, A.S.; Sebbenn, A.M.; Clement, C.; Figueira, A. 2007. High levels of genetic divergence and inbreeding in...


Received: 25/07/2016
Accepted: 12/06/2017